Integrated HPC application services for multiscale materials modeling

Ivan Kondov[1], Robert Maul[1], Stefan Bozic[1], Wolfgang Wenzel[2]
[1]Steinbuch Centre for Computing, [2]Institute of Nanotechnology
Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1,
76344 Eggenstein-Leopoldshafen, Germany
ivan.kondov@kit.edu

Computational materials science is essential for development of products with novel properties (examples: catalysis, semiconductors, and alloys). Nevertheless, next generation materials, especially nano-structured materials, exhibit pronounced complexity and multiscale behavior. Also “classical” systems like the biological cell and its building blocks include structures on many length scales. Thus, the physical models and the simulation protocols employ many different well established methods and different codes to treat the steps in these protocols. However, the lack of integration of these individual codes, the increasing complexity of models and the high demand for distributed HPC resources reduces industrial usability of the methods. In addition, joint effort of groups providing expertise for all different methods is needed. These aspects are treated within the project MMM@HPC [1] which brings together scientists from industry and academia into a unified community which is able to use the e-infrastructure to solve modern real-life problems.

We will give an overview of our methodology to provide adequate solutions for the following requirements:

Reusable interfaces and workflows. To this end, we adopted GridBeans [2] which is a modern technology to create application interfaces for use in grid middleware, such as UNICORE 6 [3] or Globus. Every GridBean provides a graphical user interface and can be readily included into different workflows without further modifications. Moreover, the workflows created for one specific task can be reused in other simulations with minor parametric modifications.

Robust tools and standards for data exchange between individual codes. In the field of materials modeling a variety of data formats are used and virtually every individual code has its own non-standard input and output formats. Thus, we aim to enhance data interoperability of individual GridBeans employing the Chemical Markup Language (CML) standard and work together with experienced developers from other projects, currently from UNICORE and OpenMolGrid [4, 5].

Solutions for licensing issues. Unfortunately, many of the codes used in the community are provided under non-free (proprietary) licenses. To treat this aspect we are working on a solution based on the Virtual Organization Membership Service (VOMS) and UNICORE.

Security and reliability. Industrial applications need secure handling (communication, storage) of simulation data. Moreover, the storage, the connection and the processes must be error-tolerant and thus reliable. All these are implemented properties of the underlying generic e-infrastructures and they are reused in our approach.

Capacity (high throughput) and capability (high performance) computing. The applications addressed in the community are particularly demanding with regard to computing and storage resources. This is why the scalable deployment of the application services requires linking to HPC and distributed resources, e.g. such as those provided in the projects EGI and PRACE.

All these user requirements pose a great challenge for both code developers and providers of e-infrastructures. These aspects are addressed in the EU project MMM@HPC. The application protocols are mapped onto scientific workflows and the application interfaces are able to ex-change input/output

data using data formats like CML. The platform of our choice is the UNICORE middleware that is broadly and productively deployed in different grid infrastructures such as D-Grid, DEISA and PRACE.

As an example, we present in Figure 1 a model of an Organic Light Emitting Diode (OLED) that requires treatment on different size scales using different code types – quantum mechanics, molecular mechanics, kinetic Monte Carlo (coarse-grained method) and finite element analysis (continuous method). Two of the interfaces are shown on the figure. We aim genericity of the developed workflow so that we consider several different codes that can perform one specific step in the simulation protocol. For example, an electronics structure calculation can be carried out employing two or more alternative programs. Our code selection criteria are maturity, open accessibility, high parallel performance, and availability of expertise by partners.

To demonstrate the functionalities of the developed tools we consider key applications, including de-novo modeling and simulation of whole devices, such as organic electronics, molecular electronics, carbon nano-device and Li-Ion batteries.

References

[1] Project MMM@HPC: Multiscale Materials Modelling on High Performance Computer Architec-tures, Available online under www.multiscale-modelling.eu.

[2] R. Ratering, A. Lukichev, M. Riedel, D. Mallmann, A. Vanni, C. Cacciari, S. Lanzarini, K. Benedyczak, M. Borcz, R. Kluszcynski, P. Bala, G. Ohme, "GridBeans: Support e-Science and Grid Applications," Proceedings of the Second IEEE International Conference on e-Science and Grid Computing (e-Science'06), p. 45, IEEE 2006.

[3] UNICORE 6 - Recent and Future Advancements, A. Streit, P. Bala, A. Beck-Ratzka, K. Benedyczak, S. Bergmann, R. Breu, J. M. Daivandy, B. Demuth, A. Eifer, A. Giesler, B. Hagemeier, S. Holl, V. Huber, N. Lamla, D. Mallmann, A. S. Memon, M. S. Memon, M. Rambadt, M. Riedel, M. Romberg, B. Schuller, T. Schlauch, A. Schreiber, T. Soddemann, W. Ziegler, Report Nr. Jül-4319, ISSN 0944-2952, FZ Jülich 2010.

[4] Sulev Sild, Uko Maran, Mathilde Romberg, Bernd Schuller, Emilio Benfenati, OpenMolGRID: Using Automated Workflows in GRID Computing Environment. In Advances in Grid Computing - EGC 2005, P. M. A. Sloot et al (Eds), Lecture Notes in Computers Sciences, Vol. 3470, pp464-473, Springer 2005.

[5] Sulev Sild, Uko Maran, Andre Lomaka, Mati Karelson, Open Computing Grid for Molecular Science and Engineering, J. Chem. Inf. Model., 46, pp953-959, 2006.

Figures

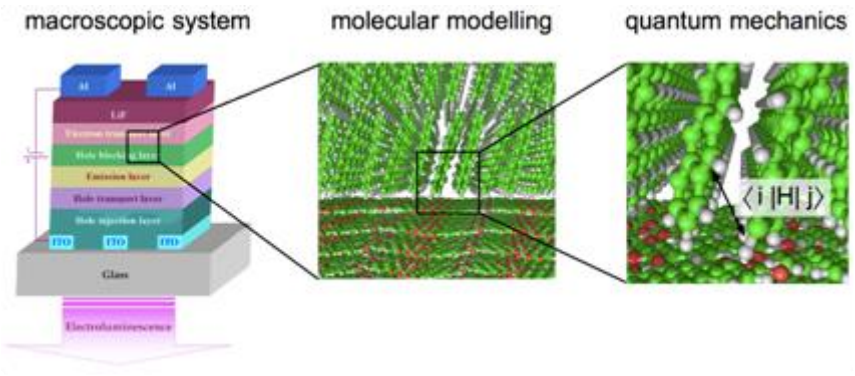

Figure1: Multiscale modeling of an Organic Light Emitting Diode (OLED).